

# 基于映射模型的风机控制器异构数据解析

吴海列<sup>1</sup>, 盛豪杰<sup>2</sup>, 邬惠峰<sup>2</sup>

(1. 浙江运达风电股份有限公司 风力发电系统国家重点实验室, 杭州 310013; 2. 杭州电子科技大学, 杭州 310018)

**摘要:**近几年来风力发电技术已成为新能源技术的重要组成部分,不同的风电公司已经相继在国内各个省份、市区部署了大批量的发电基站。在风机数量爆炸性增长的同时,产生并传递了大量的监控数据,维护人员需要定期对这些数据进行操作查看,然而数据的不统一性和复杂性为这项工作的开展造成了很大的困难。为此,创新性地构建了一种基于映射模型的异构数据解析方法。风机控制器(Beckhoff, Mita等)获取到的不同类型的监控数据,经过数据输入接口传输到映射模型,映射模型里定义多个映射约束、映射条件和映射规则,能够将传入的数据整合到同一种数据结构下,然后发送到与输出接口对接的系统软件平台,最后经过简单的图像处理在平台上显示。试验结果表明,使用这种方法大大提高了开发维护人员日常筛查数据的速度和效率。

**关键词:**映射模型;映射规则;异构数据;控制器;软件平台

**中图分类号:** TM 614      **文献标志码:** A      **文章编号:** 1674-1951(2018)04-0001-04

## 0 引言

风力发电作为新时代的一种新能源技术,具有重要的战略性地位。自20世纪90年代以来,迫于能源危机,许多国家将新能源和可再生能源的研究、开发和利用工作摆在了很重要的位置。地大物博的中国,风能资源更为丰富,据估计,平均海拔10 m范围内理论风能储量为3 226 GW,而陆上和海上的可开发储量就有1 000 GW。目前,风电行业在全国迅速发展,各种风力发电技术应运而生<sup>[1]</sup>:陈雷、邢作霞等人阐述了大型并网型风力发电机组的功率调节方式<sup>[2]</sup>;王志新、张华强等人综述了风力发电技术态势,列举了机组状态监测系统、测量参数、系统构成、实施功能等技术,还介绍了风电机组功率控制策略,创新性地提出了风电机组低电压穿越技术<sup>[3]</sup>;张新房、徐大平等人就风力发电技术在提高机组容量、改进功率调节、匀速运行、发电机和电力电子技术等方面的发展展开了论述<sup>[4]</sup>;周双喜、王海超等人对风力发电的实际效益进行了分析,并给出了风电运行价值的详细计算表达式<sup>[5]</sup>。

高数量的风场,需要耗费大量的人力、物力、财力进行维护,但是风场所在位置大多偏僻,不方便维护人员频繁到现场对风机进行故障排查和解决,这其中有些故障只是小问题,开发人员看下数据就能查明。然而,不同类型控制器的数据显示不尽相同,维护人员分析起来要针对不同的数据格式采用不同的解析方法,操作起来效率低,复杂度高。本文采取

了数据映射技术,能够将同种类型数据或者不同类型数据转换到同一种数据结构中,很好地解决了此类问题。前者在映射模型中直接转换,后者基于设定的映射规则,如基于映射表,进行间接转换。

国内外对数据映射技术有广泛地研究,比如:文献[6]提出根据一个高效和可扩展的运行库,集成对自动数据分区的映射技术;文献[7]描述了概率模块化网络和信息理论的标准数据映射,通过3个独特的学习任务将有限集合的数据点定量映射;另研究了将网络信息(拓扑,设备和状态等)映射到符合Alloy合金规范的原理和算法,以关系形式对网络信息建模,然后根据新颖的数据映射原理和算法将关系数据映射到符合Alloy合金规范的格式中<sup>[8]</sup>;Guo D.和Zhu X.等人提出了一种基于流映射的方法,从大量的移动数据中提取固有模式,并构建数据的有效视觉表示,以便了解复杂的流动趋势、消除杂散数据方差,以控制群体规范化流量,检测现有方法不能识别的高级模式<sup>[9]</sup>;此外文献[10]中,还对数据映射和常见原因故障概率进行了预测,通过不同大小的其他系统的多个故障事件数据来估计依赖系统的故障概率。

数据映射技术应用于各个领域,然而专门针对风机控制器所产生的异构数据的映射技术却少有涉及。本文提出了一种异构数据解析方法,创新性地将不同风机控制器融入其中,使其能够对获取到的不同类型数据,通过映射模型,解析成同一种数据结构,并产生直观的数据折线图,以方便维护人员查看,降低排查难度,提高排查的速度和效率。

## 1 异构数据处理系统框架

异构数据的总处理框架如图 1 所示,将解析引擎(映射模型)嵌入到软件平台中,以供内部处理数据所用,不对外开放接口;而软件平台所提供的几个主要接口是:①获取风机数据接口(单向);②与本地存储进行交互通信接口(双向);③为图像处理工具 teechart 输送数据的传送接口(单向)。

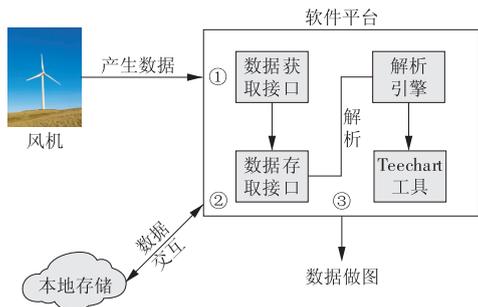


图 1 异构数据处理系统框架

### 1.1 映射模型

本节形式化描述了映射模型的几个主要数据维度和定义方法。

定义 1: 每一条获取到的监控数据  $F = \{A, T, TP, DT, LT\}$ , 分别表示为通道、值、控制器类型、时间、日志类型。

①  $A = \{Name 1, Name 2, \dots, Name k\}$ , 表示一系列通道名称的集合; 令  $dom(Name i) (i \in [1, k])$  表示通道域  $A$ 。则总通道集合  $dom(Name) = dom(Name 1) \times dom(Name 2) \times \dots \times dom(Name k)$ 。对于通道  $A$ ,  $dom(Name)$  是  $A$  的域。我们考虑的域是在大多数关系数据库中能找到的典型域, 例如整数、字符串、实数、布尔等。除了典型域以外, 我们开扩展了如下域: 时间, 空域等。

②  $T = \{valuename 1, valuename 2, \dots, valuename k\}$ , 表示为所有通道名称的值的集合。

③ 我们使用  $t[Z]$ ,  $Z$  为  $F$  中的任意有限数量子集, 表示对应属性的值。例如, 对于  $A$  的通道中的通道  $name i$  对应的值  $valuename i$  可以用  $t[valuename i]$  表示, 即  $T[valuename i]$ 。

④ 我们使用标准关系代数算子, 如映射用  $\pi X$  表示; 选择用  $\sigma X$  表示。

⑤  $TP = \{Beckhoff, Mita \dots\}$ , 表示所有风机主控制器类型集合;

⑥  $DT = \{datetimenam e 1, datetimenam e 2, \dots, datetimenam e k\}$ , 我们使用  $dt[name i]$  表示集合  $A$  中通道  $name i$  被触发时对应的日期和时间, 即  $datetimenam e i = dt[name i]$ 。

⑦  $LT = \{5MinLog, TripLog, StatusLog\}$ , 表示为

所有要解析的日志类型集合。

要实现从  $F \rightarrow F'$  的解析过程, 需要将  $F$  中每个元素  $(A, T, TP, DT, LT)$  以某种规则映射到  $F'$  中对应的元素  $(A', T', TP', DT', LT')$ , 给出以下形式化定义。

定义 2: 给定一个通道  $U$ , 对于每个  $U \in A$ ,  $t$  代表  $U$  在域  $A$  上的映射结果。  $t[A[u]]$  可以是  $dom(A)$  里的一个常数, 也可以是一个属于  $V$  的变量或者表达式  $v - S$ , 其中  $S$  是  $dom(A)$  的有限子集,  $v \in V$ ,  $V$  是一系列变量的集合,  $V \cap S = \emptyset$ 。

映射的唯一性: 为了描述一系列映射, 我们使用了映射模型代替了简单的映射关系集合。此外, 我们规定了每个变量出现在一个映射限制里, 即映射结果表中的两个不同的映射是完全独立的。

定义 3: 假设  $X, Y$  是不相交的非空通道集, 从  $X$  到  $Y$  的映射是  $X \cup Y$  上的有限映射, 使得每个变量在映射模型中最多出现一次映射。

定义 4: 映射模型中的映射结果表用  $m$  表示,  $P$  表示映射表中的某个映射关系。

### 1.2 映射约束/条件

并不是每个控制器里的任意一条数据都可以映射到同一个数据结构中显示, 所进行映射的数据必须满足以下条件和约束。

映射约束: 我们分别考虑基于映射结果表  $m$  的映射表达式  $X \rightarrow Y$  的一组映射关系  $p$  和  $p'$ , 其中  $X \in A, Y \in A$ 。  $p''$  是映射关系  $p$  和  $p'$  的笛卡尔乘积, 其中  $p$  的每个元组  $t$  与  $p'$  的每个元组  $t'$  相关联。 给定从  $X$  到  $Y$  的映射结果表  $m$ , 我们使用  $m$  作为过滤上述笛卡尔乘积条件, 即筛选所有可能映射关系。 只有满足条件  $t''[X] \in \pi X(p(m))$  和  $t''[Y] \in \pi Y(\sigma X = t''[X](p(m)))$  的映射  $p''$  才是符合映射表达式  $X \rightarrow Y$  的映射关系。

定义  $Y_m(x)$  表示通道  $X$  到  $Y$  映射得到的所有结果集, 其中  $x \in dom(X)$ ,  $Y_m(x) = \{y | \exists t \in m \& \exists p \in m, p(t[x]) = t[y] \& p(t[y]) = t[x]\}$ 。 我们规定: 假设  $U = X \cup Y$ , 即任何一个基于映射结果表  $m$  的映射表达式  $X \rightarrow Y$  的映射关系都属于  $U$ , 如果  $t[Y] \in Y_m(t[X])$ , 则称  $Y_m(x)$  为属性  $X$  基于映射表  $M$  到  $Y$  的结果集。

映射条件 1: 假设  $\exists DT_x \in A_x, DT_y \in A_y$ , 要产生基于映射结果表  $m$  的映射表达式  $X \rightarrow Y$ , 必须满足  $t[DT_x] = t[DT_y]$ 。

映射条件 2: 假设  $\exists LT_x \in A_x, LT_y \in A_y$ , 要产生表  $m$  的映射表达式  $X \rightarrow Y$ , 必须满足  $t[LT_x] = t[LT_y]$ 。

映射条件 3: 只有同时满足映射条件 1 和映射条件 2, 才能产生基于映射结果表  $m$  的映射表达式  $X \rightarrow Y$ 。

### 1.3 映射算法

异构数据映射代码见图 2 所示。

```

00 Algorithm Heterogeneous DataMap
01 Input: F fx, F fy
02 Output: F fz
03 Begin:
04   fx = do Analysis on x.csv;
05   fy = do Analysis on y.csv;
06   fz = do Cartesian on (fx, fy);
07   while (every item in fz)
08   {
09     Judge if t[DTx] = t[DTy];
10     if not:
11       fz.drop(current item);
12       next cycle;
13     Judge if t[LTx] = t[LTy];
14     if not:
15       fz.drop(current item);
16       next cycle;
17     Judge t'' [x] ∈ πx(p(m)) AND
18     Judge t'' [y] ∈ πy(σ x=t'' [x](p(m)))
19     if not:
20       fz.drop(current item);
21       next cycle;
22   }
23 End;
```

图 2 异构数据映射代码示意

在下面一个实例中,表 1 包含两个表项 F<sub>1</sub>, F<sub>2</sub>, 表 2 也包含两个表项 F<sub>3</sub>, F<sub>4</sub>; 两表做笛卡尔积,产生临时映射表 3, 包含 4 个表项; 经过映射约束和映射条件 1, 2, 3 的筛选, 最终映射结果表 4 中只剩下一个正确的表项 F', 完成从 F → F' 解析过程。

## 2 实例展示

本文的软件平台架构利用 C# winform 实现, 其调用了各个控制器的 dll 连接库, 能够远程连接风机

控制器, 从而获取监控数据, 获取到的数据文件是 csv 格式的; 在平台架构里嵌入了解析引擎, 能够对异构数据进行分析、处理; 最后利用 teechart 工具将解析出来的数据进行做图处理并显示。

如图 3 所示, 在图左侧里各个模块的分区显示的含义如下:

- ① 需要解析的日志类型显示;
- ② 日志产生的时间段筛选;
- ③ 需要异构同步的控制器类型选择;
- ④ 显示所有当前筛选条件下的 csv 文件。

图 3 右侧显示的是需要进行作图处理的数据通道, 这些数据通道里的数据是已经经过映射模型处理过的, 即已将所有能够映射到同一个数据结构下的异构数据类型全部整合到了一起。

图 4 是经过 Teechart 作图处理后的折线图, 图中主要显示了两种控制器下解析的 5 个通道(液压压

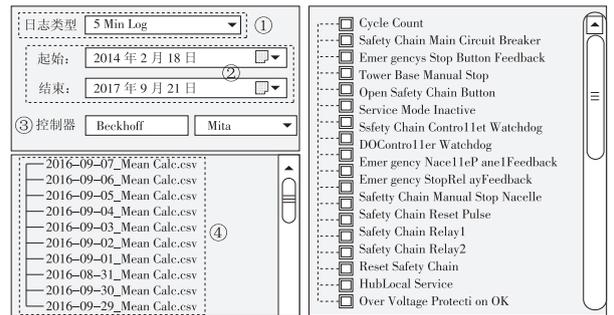


图 3 需要解析的控制器类型、日志类型、时间和通道

表 1 F<sub>x</sub> 包含的监控数据表项

表项	通道 A	值 t[A]	控制器类型 TYPE	时间 DT	日志类型 LT
F <sub>1</sub>	A <sub>1</sub>	M <sub>1</sub>	Beckhoff	T <sub>1</sub>	TripLog
F <sub>2</sub>	A <sub>2</sub>	M <sub>2</sub>	Beckhoff	T <sub>2</sub>	TripLog

表 2 F<sub>y</sub> 包含的监控数据表项

表项	通道 A	值 t[A]	控制器类型 TYPE	时间 DT	日志类型 LT
F <sub>3</sub>	A <sub>3</sub>	M <sub>3</sub>	Mita	T <sub>1</sub>	TripLog
F <sub>4</sub>	A <sub>4</sub>	M <sub>2</sub>	Beckhoff	T <sub>2</sub>	StatusLog

表 3 F<sub>x</sub> × F<sub>y</sub> 所得所有的监控数据表项

表项	通道 A <sub>x</sub>	值 t[A <sub>x</sub> ]	时间 DT <sub>x</sub>	日志类型 LT <sub>x</sub>	通道 A <sub>y</sub>	值 t[A <sub>y</sub> ]	时间 DT <sub>y</sub>	日志类型 LT <sub>y</sub>
F <sub>13</sub>	A <sub>1</sub>	M <sub>1</sub>	T <sub>1</sub>	TripLog	A <sub>3</sub>	M <sub>3</sub>	T <sub>1</sub>	TripLog
F <sub>14</sub>	A <sub>1</sub>	M <sub>1</sub>	T <sub>1</sub>	TripLog	A <sub>4</sub>	M <sub>2</sub>	T <sub>2</sub>	StatusLog
F <sub>23</sub>	A <sub>2</sub>	M <sub>2</sub>	T <sub>2</sub>	TripLog	A <sub>3</sub>	M <sub>3</sub>	T <sub>1</sub>	TripLog
F <sub>24</sub>	A <sub>2</sub>	M <sub>2</sub>	T <sub>2</sub>	TripLog	A <sub>4</sub>	M <sub>2</sub>	T <sub>2</sub>	StatusLog

表 4 筛选过的监控数据表项

表项	时间 DT	日志类型 LT	通道 A <sub>x</sub>	值 t[A <sub>x</sub> ]	控制器类型 TYPE <sub>x</sub>	通道 A <sub>y</sub>	值 t[A <sub>y</sub> ]	控制器类型 TYPE <sub>y</sub>
F'	T <sub>1</sub>	TripLog	A <sub>1</sub>	M <sub>1</sub>	Beckhoff	A <sub>3</sub>	M <sub>3</sub>	Mita

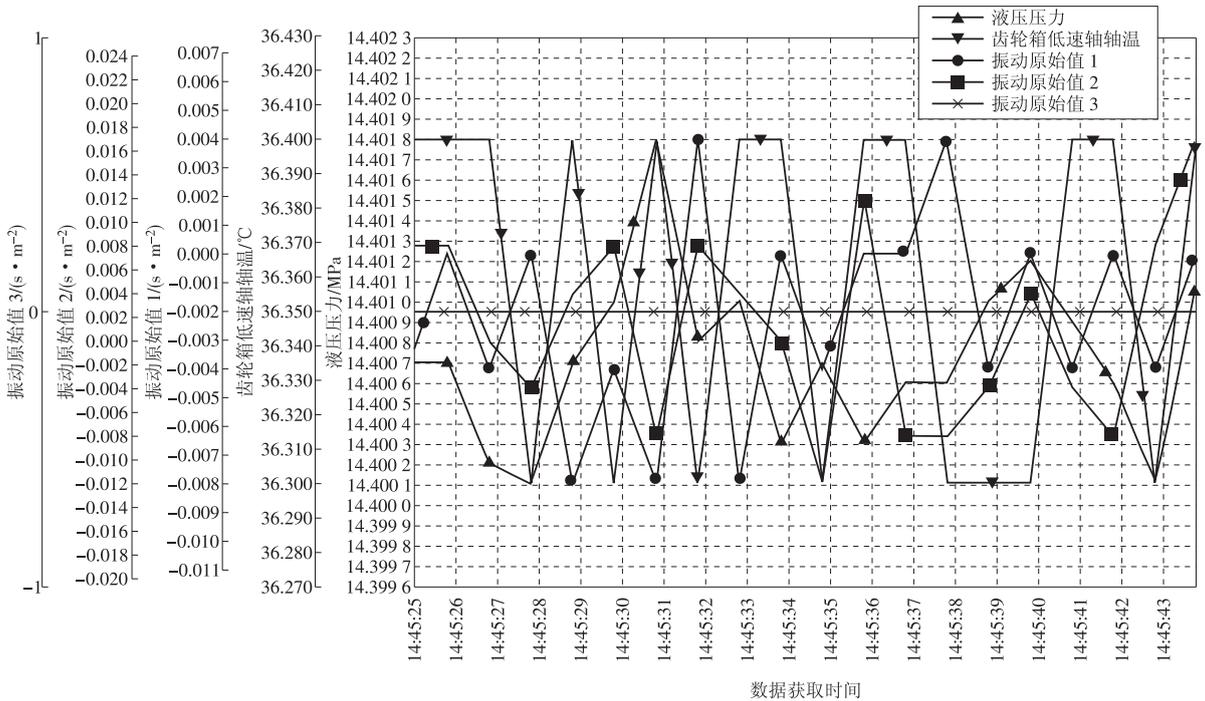


图 4 映射处理后的数据显示

力、齿轮箱低速轴轴温、振动原始值 1、振动原始值 2、振动原始值 3) 的数据结果,可以直观地进行查看和对比。

### 3 结论和展望

本文提出的映射模型能够有效处理不同风机控制器产生的数据。这些数据被映射到同一种数据结构下,然后利用软件平台的 Teechart 工具处理得到解析数据,产生方便开发维护人员查看的直观图。此外,还对映射模型的形式化定义进行了详细地描述,同时提出了需要的映射约束和映射条件,基于这些约束和映射条件,阐述了映射算法的主要思想。

在未来的工作中,主要针对数据处理的时间和效率提升进行研究。由于风机的数量庞大,产生的数据参差不齐,种类繁多,当前的数据映射模型只考虑了映射结构统一的问题,处理速度和效率上并没有深入研究。下一步的任务是在原有映射模型基础上,进一步改进映射结构,提高数据映射速度和准确性。

#### 参考文献:

[1] WANG Xiaorong, WANG Weisheng, DAI Hui. Wind power industry in China[J]. Electricity, 2004(1): 41 - 43.

[2] 陈雷,邢作霞,潘建,等. 大型风力发电机组技术发展趋势[J]. 可再生能源, 2003(1): 27 - 30.

[3] 王志新,张华强. 风力发电及其控制技术新进展[J]. 通用低压电器, 2009(19).

[4] 张新房,徐大平,吕跃刚,等. 风力发电技术的发展及若干问题[J]. 现代电力, 2003, 20(5): 29 - 34.

[5] 周双喜,王海超,陈寿孙. 风力发电运行价值分析[J]. 电网技术, 2006, 30(14): 8 - 102.

[6] GONZALEZ-Escribano A, TORRES Y, FRESNO J, et al. An extensible system for multilevel automatic data partition and mapping[J]. IEEE Transactions on parallel and distributed systems, 2014, 25(5): 1145 - 1154.

[7] WANG Y, LIN S, LI H, et al. Data mapping by probabilistic modular networks and information theoretic criteria[J]. IEEE Transactions on signal processing, 1998, 46(12): 3378 - 3397.

[8] LI F, AN C, YANG J, et al. Data mapping principles and algorithm for self-configurable systems[J]. China communication, 2015, 12(10): 169 - 181.

[9] GUO D, ZHU X. Origin-destination flow data smoothing and mapping[J]. IEEE Transactions on visualization and computer graphics, 2014, 20(12): 2043 - 2052.

[10] XIE L, ZHOU J, WANG X. Data mapping and the prediction of common cause failure probability[J]. IEEE Transactions on reliability, 2005, 54(2): 291 - 296.

( 本文责编:陆华)

#### 作者简介:

吴海列(1979—),男,浙江宁波人,工程师,工学硕士,从事风电行业风机主控程序开发与研究工作(E-mail: wuhl@chinawindey.com)。

盛豪杰(1992—),男,浙江温州人,杭州电子科技大学工学硕士,从事 FBD 平台开发的调试研究工作(E-mail: sheng\_haojie@126.com)。

鄢惠峰(1978—),男,浙江平湖,工程师,工学博士,从事嵌入式系统、智能控制开发与研究工作(E-mail: whf@hdu.edu.cn)。